
Exploración de Redes Sociales en la Producción Académica Estadística

Exploring Social Networks in Academic Statistical Production

Ariadna Contreras^a
arcontrerasa@unal.edu.co

Juan Sosa^b
jcsosam@unal.edu.co

Resumen

Este trabajo desarrolla un análisis de redes sociales a los artículos publicados por la Revista Colombiana de Estadística, esto con el fin de descubrir patrones y características importantes en las relaciones de citas tanto dentro de los autores que pasaron o fueron referenciados por la revista como dentro de los artículos enlazados a estos autores. Mediante un análisis descriptivo y una partición de las redes de interés se descubrió el papel fundamental que juegan no solo autores reconocidos mundialmente, sino también docentes de la Universidad Nacional de Colombia que han dedicado su vida al campo de la estadística.

Palabras clave: Centralidad; grafos; redes sociales; Revista Colombiana de Estadística.

Abstract

This work develops a social network analysis of the articles published by the Revista Colombiana de Estadística, with the aim of uncovering important patterns and characteristics in citation relationships, both among the authors who have been cited or referenced by the journal and among the articles linked to these authors. Through a descriptive analysis and a partitioning of the networks of interest, the fundamental role played not only by internationally recognized authors but also by professors from the National University of Colombia who have dedicated their lives to the field of statistics was discovered.

Keywords: Centrality; graphs; social networks; Colombian Journal of Statistics.

^aDepartamento de Estadística, Universidad Nacional de Colombia

^bDepartamento de Estadística, Universidad Nacional de Colombia

1. Introducción

En la era actual, donde la cantidad de información poco se considera un impedimento para analizar una población, el análisis de las redes se ha convertido en una herramienta fundamental para comprender la dinámica y el impacto de las relaciones existentes en un conjunto de objetos; es el caso del internet, que conecta a millones de computadoras y dispositivos móviles alrededor del mundo, permitiendo el intercambio de información. Otro ejemplo de la vida cotidiana son los sistemas de emergencias, estos permiten que diferentes hospitales, estaciones de policía y bomberos estén conectados entre sí para poder brindar un mejor servicio de salud o seguridad a una comunidad.

Este trabajo se enfoca en un ámbito específico: el análisis de las redes de citas en los artículos publicados por la *Revista Colombiana de Estadística* (<https://revistas.unal.edu.co/index.php/estad>), dirigida por el Departamento de Estadística de la Universidad Nacional de Colombia. Esta revista, reconocida por su contribución en el campo de la estadística y su relevancia dentro del contexto académico colombiano, ha publicado más de 300 artículos desde 1969 sobre la aplicación, educación, investigación e historia en estadística. Ofreciendo una amplia cantidad de recursos (artículos) que pueden ser utilizados en diversos campos como lo son el lenguaje natural o el análisis de redes sociales. En este último caso se consideran como objetos tanto los autores como los artículos, y la relación definida entre cada uno de estos subgrupos son las citas, obteniendo una red de citas entre autores y una red de citas entre artículos, las cuales reflejan las tendencias, influencias y evolución del pensamiento estadístico a partir de la década de los sesenta a la fecha en Colombia.

Entre los múltiples estudios sobre el área del análisis de redes, se encuentran los artículos de Maltseva and Batagelj (2021) y Otte and Rousseau (2002), en los cuales se realiza un análisis a las citas realizadas entre un conjunto de artículos de diversas temáticas. En ambos casos se utilizaron artículos cuya temática fuera el análisis de redes sociales, sin embargo, para este trabajo se desea ampliar el conjunto de artículos a múltiples áreas de la estadística, el objetivo de este estudio se centra primero en visualizar la estructura de la red de citas, identificando los trabajos más influyentes según diferentes métricas locales y los patrones de colaboración entre autores a lo largo del tiempo mediante métricas globales. Segundo, aportar para el desarrollo de las redes de citas en la difusión del conocimiento estadístico tanto a nivel nacional como internacional, proporcionando una visión única de la interconexión académica y su evolución en el tiempo. Para lograr estos objetivos se emplean metodologías descriptivas del análisis de redes sociales, las cuales incluyen, pero no se limitan, al análisis de centralidad para identificar artículos y autores clave. Esto con el fin de describir el comportamiento general de la red y vislumbrar comunidades que se pueden generar dentro de esta; sin embargo, vale la pena aclarar que los resultados obtenidos están basados en la calidad de la información extraída de los artículos, por lo cual teniendo en cuenta que la exigencia en la redacción de las citas a lo largo de los años ha ido mejorando, hay información que se encuentra incompleta en las décadas de los 60

a los 80.

Este trabajo no solo aporta a la comprensión de la dinámica de las citas en el campo de la estadística en Colombia, sino que también ofrece los primeros pasos para tener una perspectiva valiosa sobre cómo las comunidades académicas se forman y evolucionan, y cómo el conocimiento se construye y se expande en el ámbito científico. Al hacerlo se espera empezar a contribuir a la literatura existente sobre análisis de redes en la investigación científica y proporcionar una herramienta útil para futuras investigaciones en el campo de la estadística y más allá.

2. Conceptos básicos de redes sociales

Una **red** se define como una colección de objetos interconectados (Sosa (2022)), la cual se puede visualizar por medio de un **grafo** G , este se compone de un conjunto de objetos V (nodos o vértices) y un conjunto de enlaces E (aristas) que conectan los nodos, permitiendo que un grafo se defina como $G = (V, E)$. La Figura 1, muestra un ejemplo de grafo donde el conjunto de nodos tiene 7 elementos y el conjunto de aristas tiene 42 elementos.

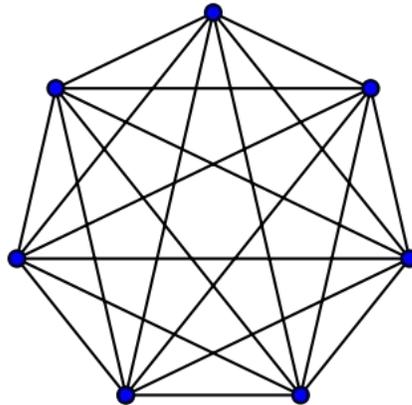


Figura 1: Visualización de un grafo, recuperado de Wikipedia (2020).

Estas representaciones matemáticas permiten analizar las interacciones entre los elementos del conjunto de nodos, conduciendo a conclusiones sobre su comportamiento colectivo y sus propiedades a nivel global. De la misma manera, un **subgrafo** de $G = (V, E)$ se define como $G' = (V', E')$ si $V' \subset V$ y $E' \subset E$, es decir, se considera un subconjunto de los nodos de G , con las respectivas aristas entre ellos.

2.1. Tipos de relaciones

Un grafo es una representación abstracta que permite modelar conexiones entre diferentes entidades, cuyo análisis se basa en las relaciones que describen estas conexiones. Estas relaciones pueden variar dependiendo de la naturaleza del grafo, lo que da lugar a clasificaciones que reflejan diferentes formas de interacción o conexión entre los elementos. Entre los aspectos más relevantes para clasificar las relaciones dentro de un grafo se encuentran su dirección y la posibilidad de asignar valores o pesos a las conexiones. Estos tipos de relaciones son fundamentales para entender y analizar las propiedades de los grafos y su aplicabilidad en diversas áreas, como redes sociales, sistemas de transporte y modelos biológicos.

2.1.1. Redes dirigidas y no dirigidas

Una relación **no dirigida** (simétrica) es aquella que no tiene dirección, es decir, para $e \in E$ se tiene $e = \{u, v\} = \{v, u\}$ con $u, v \in V$, mientras que una relación **dirigida** (no simétrica) es aquella para la cual importa la dirección de la relación, es decir, para $e, e^* \in E$ se tiene $e = \{u, v\} \neq \{v, u\} = e^*$ con $u, v \in V$. Por lo que una red se denomina dirigida si todas sus relaciones son no dirigidas, análogamente la red se denomina no dirigida si todos sus enlaces son no dirigidos. En la Figura 2 se pueden observar ejemplos de cada uno de estos tipos de redes.

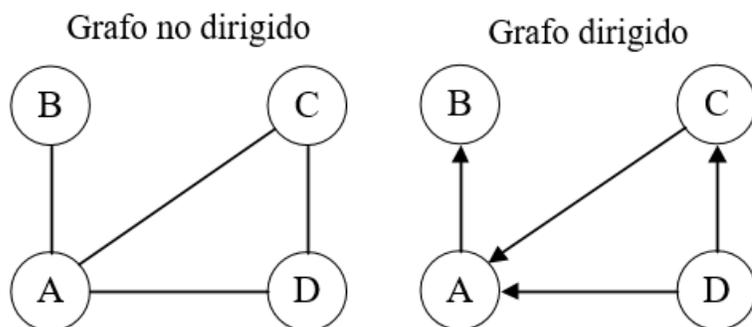


Figura 2: Ejemplo de grafo dirigido y no dirigido, recuperado de Bianco (2016)

2.1.2. Redes binarias y ponderadas

Una relación **binaria** (dicotómica) es aquella que solo asume dos valores, ausencia o presencia de la relación, mientras que una relación **ponderada** (numérica) puede tomar más de dos valores, caracterizando la relación entre dos nodos. Por lo que una red se denomina binaria si todos sus enlaces también lo son, análogamente una red se denomina ponderada si todos sus enlaces están caracterizados por un valor (peso).

2.2. Matriz de adyacencia

Otra manera de definir una red es mediante su **matriz de adyacencia**, la cual es de tamaño $|V| \times |V|$ y en cada posición se tiene el peso de la conexión entre los respectivos nodos. Como se puede observar en la Figura 3, dependiendo el tipo de red con la que se trabaje, esta matriz va a tener unas características específicas. Como se puede ver en la imagen de la izquierda, la matriz de una red dirigida no necesariamente es simétrica, caso contrario al de la matriz de una red no dirigida, mientras que gracias a la figura de la derecha se observa que al considerar una red ponderada los elementos de la matriz dejan de ser solamente 0 y 1, debido a que en estas posiciones se considera el peso de cada conexión en cuestión.

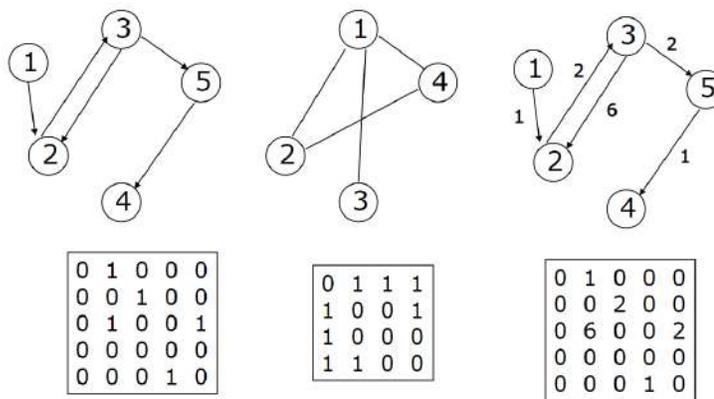


Figura 3: Matriz de adyacencia para diferentes redes, recuperado de Muñiz R. (2015).

2.3. Medidas descriptivas

El análisis de un grafo incluye la utilización de medidas descriptivas que permiten caracterizar su estructura y las relaciones entre sus elementos. Estas medidas pueden ser locales, enfocándose en propiedades específicas de los nodos o sus conexiones inmediatas, o globales, abordando aspectos que describen el comportamiento y la organización del grafo en su totalidad. Entre las medidas locales destacan aquellas relacionadas con las interacciones directas de un nodo, mientras que las globales evalúan características como la conectividad, cohesión y patrones de interacción en el grafo completo. Estas herramientas son esenciales para comprender la dinámica y el funcionamiento de sistemas complejos modelados mediante grafos, y su estudio ofrece una visión profunda sobre la estructura de las redes.

2.3.1. Medidas locales

Una **medida** se llama medida local cuando está definida para cada nodo o para cada arista de la red, es decir, que no describe el comportamiento general, sino el comportamiento específico de cada nodo. Algunas de estas medidas se describen a continuación.

- **Grado:** se entiende como el número de conexiones que tiene un nodo, en el caso de tener una red dirigida esta medida se puede calcular considerando las aristas de entrada o de salida para obtener el grado de entrada o el grado de salida respectivamente. Esta medida se define como:

$$d_v = |\{\{v, u\} \in E : u \in V\}|.$$

- **Fuerza:** se entiende como la suma de los pesos de las conexiones incidentes en un nodo, en el caso dirigido como estos pesos siempre son 1 la fuerza concuerda con el grado. Considerando que $w_{\{v,u\}}$ es el peso de la arista entre los nodos u y v , esta medida se define como:

$$s_v = \sum_{u \in V: \{v,u\} \in E} w_{\{v,u\}}.$$

- **Centralidad:** estas medidas están diseñadas para cuantificar la ‘importancia’ de los nodos en una red, existen diferentes métodos para calcular esta ‘importancia’. Según NetworkX (2023) se definen los siguientes tipos de medidas de centralidad:

- **Centralidad de cercanía:** los nodos importantes según esta medida son aquellos para los cuales, a pesar de tener pocas conexiones, sus arcos permiten llegar a todos los nodos de la red más rápido que desde cualquier otro punto de la red. Esta medida se define como:

$$C_C(v) = \frac{n-1}{\sum_{u \in V} d(v,u)}$$

donde $d(u, v)$ es la menor distancia para llegar del nodo u al nodo v y n el número de nodos conectados a v .

- **Centralidad propia:** es una métrica de centralidad recursiva, para la cual un nodo se considera importante si al mismo tiempo sus vecinos son importantes o centrales. El vector de centralidades x se calcula como la solución a la ecuación:

$$\lambda x^T = x^T A.$$

donde A es la matriz de adyacencia de la red y λ es el valor propio más grande asociado a A .

- **PageRank:** considera el número y la calidad de los enlaces incidentes en un nodo, se puede considerar una variación de la centralidad propia,

en términos de que un nodo depende de sus vecinos, los cuales traspasan una fracción de su importancia al nodo en cuestión. Esta medida dado Page (2001) se define como:

$$PR(v) = (1 - d) + d \sum_{u \in V: \{u,v\} \in E} \frac{PR(u)}{C(u)}.$$

2.3.2. Medidas globales

Una métrica se llama medida global cuando resume un aspecto de la conectividad de la red, permitiendo hacer comparaciones entre redes basadas en estas métricas. A continuación, según Sosa (2022) se describen algunas de estas medidas:

- **Densidad:** Se define como la frecuencia relativa de las aristas observadas respecto al potencial de aristas en la red. Esta se calcula de la siguiente manera:

$$Den(G) = \frac{|E|}{|V|(|V| - 1)/2}.$$

- **Transitividad / Coeficiente de agrupamiento:** Se entiende como la propensión de una red a formar triángulos en sus conexiones. Considerando como una **tripla** aquella que está constituida por 3 nodos que están conectados por 2 o 3 aristas, y un **triángulo** como 3 nodos totalmente conectados que tienen una forma de triángulo siguiendo o no la dirección de sus aristas, para el caso dirigido y no dirigido respectivamente. Entonces la medida se define como:

$$Tr(G) = \frac{3 \times \#\text{triángulos}}{\#\text{triplas}}.$$

- **Reciprocidad:** Se entiende como la tendencia con la que hay reciprocidad entre las aristas que conforman la red. Esta se define como:

$$Rec(G) = \frac{\#\text{aristas recíprocas}}{\#\text{aristas}}.$$

2.4. Agrupamiento

La **partición de redes** también conocida como **detección de comunidades** se refiere a la **segmentación** del conjunto de vértices en subconjuntos 'naturales' respecto a sus patrones relacionales. Los algoritmos de agrupamiento de grafos buscan una partición $C = \{C_1, \dots, C_k\}$ del conjunto de vértices V , de modo que el número de enlaces entre dos particiones C_k y C_l sea relativamente pequeño comparado con el número de enlaces dentro de una partición C_k . Entre estos algoritmos se encuentran los relacionados al **particionamiento espectral**, los cuales son particularmente útiles en situaciones donde la estructura de la red no

es obvia a simple vista y puede revelar subgrupos o comunidades que no son evidentes mediante otros métodos de análisis de redes. El concepto clave detrás del particionamiento espectral son los valores y vectores propios de la **matriz de modularidad**, calculada a través de la matriz de adyacencia y una matriz diagonal de los grados de los nodos. El algoritmo utilizado en este trabajo es el **algoritmo de Louvain**.

2.4.1. Algoritmo de louvain

Según Blondel et al. (2008) el algoritmo de Louvain es un método popular y eficiente para detectar comunidades en redes de gran tamaño, es particularmente conocido por su rapidez y su capacidad para encontrar comunidades de alta calidad en estas redes. Se basa en la optimización de la **modularidad**, una medida que cuantifica la densidad de enlaces dentro de las comunidades en comparación con los enlaces entre comunidades.

3. Metodología

3.1. Base de datos

Para este trabajo se utilizaron 527 artículos publicados por la Revista Colombiana de Estadística desde su creación en 1968 hasta el día de hoy. Para cada uno de estos artículos se extrajo la siguiente información:

- Autores del artículo.
- Título del artículo.
- Año de publicación.
- Década de publicación.
- Autores de los artículos a los que se hizo referencia.
- Título del artículo referenciado.
- Editorial, institución educativa, conferencia, o empresa a nombre de la cual fue publicado el artículo referenciado.
- Año de publicación del artículo referenciado.

Después de extraer la información (todos los artículos estudiados tienen disponible esta información) se realizó una limpieza de datos, esto debido a que un mismo autor puede ser referenciado de distintas maneras, mismo caso con la editorial o institución a nombre de la cual fue publicada la referencia. El resultado final de este proceso fue una base de datos relacional con 53321 registros, el cual es el número de referencias realizadas por artículos publicados por la revista.

3.2. Aplicación

Las bases de datos y el código utilizado para aplicar metodología especificada se encuentran en: <https://github.com/AriadnaAbril/>. De las variables extraídas para cada artículo se puede definir una relación de citas tanto dentro de un conjunto de autores como de artículos, por lo que se van a considerar dos tipos de redes:

- **Redes de autores:** esta clase de redes considera a los autores como nodos, y sus enlaces como la relación “cita a...” por lo que estos se definen como dirigidos. Así mismo como un autor puede citar a otro por medio de diferentes artículos, estas conexiones se van a ponderar por el número de citas que realizó un autor a otro. Una característica a resaltar es la presencia de bucles, los cuales pueden existir dentro de esta clase de redes debido a que un autor puede citarse a sí mismo.
- **Redes de artículos:** esta clase de redes considera a los artículos como nodos, y al igual que en el caso anterior la relación entre ellos es “cita a ...”, teniendo de nuevo una red dirigida. En este caso, para caracterizar la relación se considera la diferencia en años de las fechas de publicación, volviéndose una red ponderada. Por último, teniendo en cuenta la información disponible, los vértices se caracterizan tanto por el año de publicación de los artículos como por la editorial, conferencia o institución a nombre de la cual fue publicado el artículo.

Debido a la amplia ventana de tiempo que se considera en este estudio se decidió construir una red independiente para cada década y otra que unifique todas estas, esto no solo con el fin de analizar la evolución del comportamiento relacional de las citas en las publicaciones de la Revista Colombiana de Estadística sino también analizar las interacciones en el conjunto total de autores y artículos que han pasado por la revista. Para las redes que consideran la totalidad de artículos y autores en los 55 años de publicaciones se realiza una detección de comunidades, con el fin de encontrar un patrón de segmentación en los nodos según la conectividad de la red. A causa del gran tamaño de estas redes, el algoritmo que se utilizó para la identificación de comunidades fue el de Louvain, por su capacidad para optimizar la modularidad de manera eficiente, incluso en redes de gran tamaño, como las analizadas en este estudio. Este algoritmo permite identificar comunidades altamente conectadas internamente y con pocas conexiones externas, lo cual es ideal para redes académicas donde es crucial detectar grupos de autores o artículos con intereses o influencias comunes. Entre sus ventajas destacan su rapidez, necesaria para procesar grandes volúmenes de datos, y su coherencia al generar particiones naturales en la red, reflejadas en altos valores de modularidad, lo cual asegura que las comunidades identificadas no solo sean significativas, sino también representativas de las interacciones académicas subyacentes. Posterior a la detección de las comunidades se identifica la de mayor tamaño debido ya que en ella se puede identificar más fácilmente el patrón de segmentación utilizado, obteniendo una nueva red a analizar como se describe a continuación.

Cada una de estas redes fue graficada teniendo en cuenta algunas medidas locales tanto de los nodos como de las aristas, tales como:

- **Peso:** normalizando los valores de los pesos entre 0 y 1 la transparencia de las aristas representa es proporcional a estos valores, teniendo que un color más oscuro de arista representa un mayor peso.
- **Centralidad propia - PageRank:** los nombres y colores de los nodos en las redes analizadas dependen de sus valores de centralidad propia y PageRank. Los nodos con valores superiores a percentiles específicos (80, 80, 90, 95, 97, 97, 97, 90 y 99, para las décadas de los 60s a los 2020 seguido de la red que unifica estas décadas y su comunidad de mayor tamaño respectivamente) se destacan con colores diferenciados: azul oscuro para centralidad propia, rojo para PageRank y morado cuando ambos criterios se cumplen. Estos percentiles fueron seleccionados para resaltar únicamente los nodos más relevantes, asegurando una representación visual clara y evitando sobrecargar la interpretación de la red al destacar un número excesivo de nodos.

La elección de estas métricas responde a su relevancia en redes de citación. El PageRank considera tanto la cantidad como la calidad de las conexiones entrantes, lo que lo hace útil para identificar autores o artículos que ganan relevancia a través de citas por nodos destacados, incluso con pocas conexiones directas. Por su parte, la centralidad propia mide la popularidad de un nodo en función de sus conexiones directas, reflejando su importancia en el campo a pesar de ignorar la calidad de las mismas. Valores bajos de PageRank pueden sugerir una prominencia global limitada, pero no afectan significativamente el análisis si se complementan con métricas como el grado o la centralidad propia. La representación visual basada en estas métricas permite identificar y comparar la segmentación de nodos según su importancia relativa en la red, optimizando tanto la interpretabilidad como la validez del análisis.

Habiendo definido las anteriores redes, a cada una de ellas se le realizó un análisis descriptivo y un análisis de influencia dentro de las mismas. El primero se centra en describir el comportamiento general de la red dada su conectividad, esto mediante las medidas globales como número de nodos, número de aristas, grado promedio, fuerza promedio, densidad, transitividad y reciprocidad. En el caso de las redes de artículos, no se considera esta última medida debido a que es imposible que existan relaciones simétricas entre dos artículos por su componente temporal, es decir, si un artículo 1 cita a un artículo 2 es porque el 1 fue publicado después del 2, siendo imposible que el artículo 2 cite al artículo 1, razón por la cual tampoco se considera la transitividad en este tipo de redes.

Para el análisis de influencia, contrariamente al anterior análisis, se consideran medidas locales, de modo que este tipo de estadísticas permiten realizar un ordenamiento de los nodos. Las medidas que se consideran en este trabajo son la centralidad de cercanía, centralidad de intermediación, centralidad propia, *PageRank*, grado de entrada y fuerza. Como se puede ver, solo se considera el grado de

entrada y no de salida, debido a que el de entrada indica la popularidad de un nodo y el de salida la sociabilidad de este mismo. Es así, que en este contexto vale más la pena identificar a los nodos más citados debido a la importancia de su contribución al área de la estadística. En el contexto de las redes de citación, cada medida aporta una perspectiva única sobre la dinámica de las conexiones, la centralidad de cercanía identifica nodos que pueden alcanzar otros rápidamente, destacando su rol en la difusión del conocimiento, mientras que la centralidad de intermediación detecta nodos que actúan como puentes entre comunidades, cruciales para conectar áreas de investigación. Por su parte, la centralidad propia y el grado de entrada miden la popularidad de un nodo en función de sus conexiones directas, reflejando su relevancia local, aunque ignoran la calidad de estas conexiones, lo que es complementado por el PageRank, que considera tanto la cantidad como la importancia de los nodos citantes, resaltando su influencia global. Finalmente, la fuerza evalúa la intensidad de las conexiones, útil en redes ponderadas. Valores bajos en estas medidas pueden indicar un rol periférico o de menor importancia, pero no necesariamente afectan el análisis si se combinan métricas que capturen diferentes aspectos de la red. Dependiendo de los objetivos, estas métricas, incluso con valores bajos, pueden ser suficientes para identificar patrones locales o contribuciones específicas, mientras que el análisis global podría requerir mayor conectividad y centralidad.

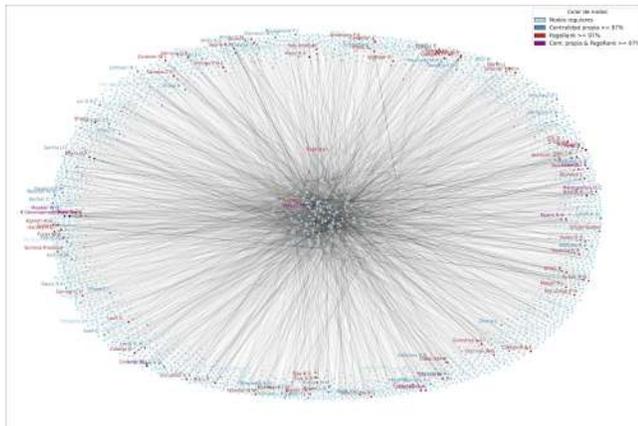
4. Resultados

red Los resultados presentan un análisis detallado de las redes de citación tanto entre autores como entre artículos. Los hallazgos incorporan diversas medidas y visualizaciones que permiten comprender la dinámica y las interacciones dentro de estas redes.

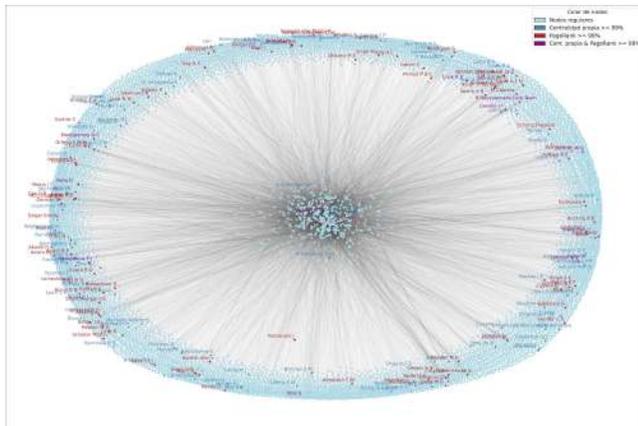
4.1. Redes de citaciones entre autores

De la representación de las redes por medio de sus grafos 4 se puede observar que conforme avanzan las décadas el número de autores importantes o centrales aumenta proporcionalmente al número de nodos. Esto es debido a que el percentil que se considera como límite para clasificar a un nodo como importante va aumentando conforme avanzan las décadas, excepto para la década de los 2020, cuando a pesar de haber un decaimiento en el número de nodos, la cantidad de autores considerados centrales es alta para el tamaño de su red.

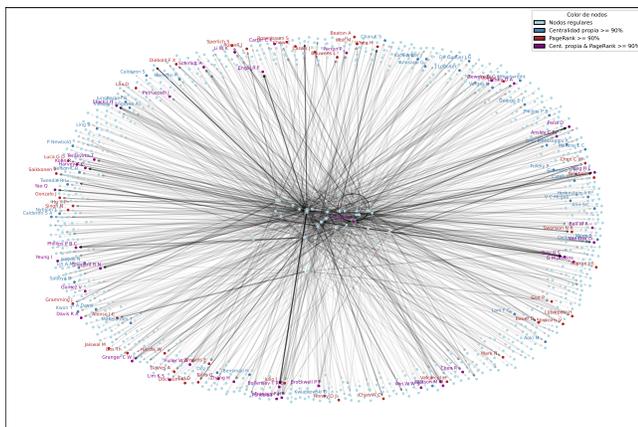
Así mismo, desde la década de los 80s se puede identificar la presencia de personajes de gran importancia en el área de la estadística tales como Cochran W G, A N Kolmogorov, Lehman E L y Tukey J W en los años mencionados, cuyas contribuciones se centran en conceptos ampliamente utilizados en diferentes campos, como lo son las técnicas de muestreo, la teoría de la probabilidad, las pruebas de hipótesis no paramétricas, o el desarrollo de métodos descriptivos como BoxPlots.



(a) Rede de citações entre autores para artigos publicados na década de 2000s



(b) Rede de citações entre autores para artigos publicados em toda a história da Revista Colombiana de Estadística



(c) Comunidade de maior tamanho para a rede de autores

Figura 4: Redes de autores.

En décadas posteriores como en los 90s entran en la discusión Neyman J y Mar-dia K V, actores conocidos por introducir el término de intervalos de confianza en pruebas de hipótesis y pruebas de normalidad multivariada respectivamente. En los años 2000 empiezan a ser reconocidos Montgomery D C, Siegel Sidney y Box G E P, siendo el último destacado como una de las mentes más brillantes del siglo XX, pero sobre todo es de destacar la presencia de los lenguajes SAS y R development, mostrando el comienzo del uso masivo en herramientas tecnológicas para el desarrollo de ideas estadísticas. Pero no solamente se destacan estadísticos reconocidos mundialmente, lo cual es de esperar, ya que por la popularidad de sus nombres sus trabajos son más tenidos en cuenta, paralelamente aparecen varios docentes del departamento de estadística de la Universidad Nacional, así como la misma institución, tales como Corzo J, Vargas J A, Blanco L, Fabio H Nieto, Diaz L G, Sosa J C, Vergara Cardozo S, Calsderon S A y Cepeda Cuervo E, donde los tres primeros corresponden a las décadas del siglo XX.

Se puede ver de las tablas 15, 16, 35 y 36, que las redes de los 60s y 70s muestran un mismo comportamiento gracias a su baja densidad y fuerza de conectividad, esto en términos de que la fuerza de un nodo es inversamente proporcional a la importancia que transmite a sus referencias, así como los autores considerados importantes según la centralidad de cercanía son aquellos que pertenecen a la componente más grande de la red, porque pueden llegar a una mayor cantidad de nodos mediante las conexiones existentes en esta, y por último debido a que la centralidad propia, también conocida como centralidad de grado, tiene en cuenta las relaciones de entrada todos los autores en la periferia de la red tienen el mismo nivel de centralidad.

Medida	Decada								
	60s	70s	80s	90s	2000	2010	2020	Completa	Comunidad
Num. Nodos	11	8	537	745	2403	6223	3278	10993	793
Num. Aristas	9	6	718	1445	6681	22068	10441	40924	1739
Grado Prom.	0.818	0.75	1.337	1.939	2.78	3.546	3.185	3.723	2.193
Fuerza Prom.	1.636364	1.500000	3.232775	4.569128	6.682480	9.495742	8.029896	9.557900	5.709962
Densidad	0.081818	0.107143	0.002495	0.002607	0.001157	0.000570	0.000972	0.000339	0.002769
Transitividad	0.000000	0.000000	0.004480	0.006019	0.000360	0.000057	0.000376	0.000315	0.001294
Reciprocidad	0.000000	0.000000	0.002786	0.001384	0.000000	0.000000	0.000000	0.000098	0.000000

Tabla 1: Medidas descriptivas globales de todas las redes de autores.

Como se observa en las tablas 17 y 37 en la década de los 80s se obtiene el mayor coeficiente de reciprocidad entre todas las redes, lo cual se considera puede suceder por dos razones: la primera es la **continuación de trabajos pasados**, es decir, cuando un autor referencia a otro puede causar un interés por el autor más antiguo en continuar su trabajo inicial a causa de nuevos descubrimientos, o como segunda razón está el **descubrimiento de autores**, dicho de otro modo, esa referenciación puede causar una introducción del autor que referencia al círculo de atención del autor original permitiendo que conozca otros de sus trabajos y sirva como base para un trabajo futuro. Otro factor a resaltar es la presencia de bucles en el conjunto de autores, característica presente también en la década de los 90s y de los 2020, los cuales aumentan el grado de entrada de los autores que realizan referencias,

estos se pueden distinguir en el centro del grafo 6(c), 7(a) y 7(c), aumentando la posibilidad de heredar importancia de sí mismos, como es el caso de Poltronieri J, O Barbary o F H Nieto respectivamente, que se encuentran en el Top 5 según el *PageRank* y/o Centralidad propia. Sin embargo, de esta época hasta la actualidad, este valor es muy cercano a 0, al igual que los valores de centralidad de cercanía, fenómeno producido por la baja densidad y grado promedio de la red, implicando pocas referencias hacia un autor y entre autores, teniendo poca contribución de los vecinos en la importancia de un autor, y al mismo tiempo aumentando las distancias mínimas necesarias para llegar de un autor a otro, causa de valores cercanos a 0 en la centralidad de cercanía de los nodos. Pero aunque no se puede hacer un análisis de criticidad en la influencia de la red, si se pueden comparar los autores de esta en términos de importancia relativa de la red, como por ejemplo en los 80s y 90s cuando debido a la gran diferencia de valores entre el primer y segundo puesto de el top se puede afirmar que Poltronieri J y Mardia K V son los autores más importantes entre los respectivos conjuntos de nodos.

A inicios del siglo XXI se destaca Montgomery a causa del gran número de autores que lo referencian, posicionándolo en el top tomando como criterio la centralidad propia, sin embargo esto no sucede considerando la centralidad de cercanía y el *PageRank*. Lo cual indica que un gran número de autores de esta época jamás se relacionaron con las áreas trabajadas por Montgomery, afirmación respaldada por 4(a), ya que este autor se encuentra en la periferia de la red junto a miles de autores para los cuales es imposible llegar a este autor a causa de la direccionalidad de la red. Simultáneamente, aquellos autores que lo referencian son considerados poco importantes, transmitiéndole una cantidad casi nula de importancia. En el caso de la década de los 2010 el autor a resaltar es Gupta R D, autor que aunque empezó a publicar hasta el 2018 es más importante que el 80% de los autores de esta red como se puede observar en 7(b). En esta misma década y la que le sigue se puede ver mucho más la importancia de herramientas computacionales, por la presencia de R Development en el top de casi todas las medidas de importancia.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	R Development Core Team	Vargas J A	R Development Core Team	R Development Core Team	L A Lopez P
2	Montgomery D C	Brockwell P J	Searle S R	Searle S R	E Castaño
3	Rubin D B	Rousseeuw P J	Schmid Friedrich	Montgomery D C	P Martinez Camblor
4	Gelfand A E	Davis R A	Schader Martin	Rubin D B	O O Melo
5	McCullagh P M	Guevara R D	Agresti Alan	Agresti Alan	J A Achcar

Tabla 2: Top autores más importantes en la red de citaciones de la década de los 00s.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	R Development Core Team	F H Nieto	R Development Core Team	R Development Core Team	V Leiva
2	Kotz S	Chen C W	Balakrishnan N	Balakrishnan N	J A Achcar
3	Johnson N L	Tong H L	Cordeiro G M	Kotz S	B Oluyede
4	Rubin D B	Peña D	Kotz S	Nadarajah S	M P Galindo
5	Efron B	Akaike H A	Singh H P	Rubin D B	R Development Core Team

Tabla 3: Top autores más importantes en la red de citaciones total entre autores.

La dinámica de citación observada en la Revista Colombiana de Estadística refleja

un interesante contraste entre el impacto local e internacional en el campo de la estadística. A nivel local, el análisis muestra que los autores afiliados a la Universidad Nacional de Colombia, como Fabio H. Nieto y Juan A. Vargas, son pilares de la producción académica nacional. Esto se refleja en su importancia dentro de las redes de citas, donde sus contribuciones actúan como nodos centrales en comunidades locales, indicando su influencia en la formación y diseminación del conocimiento estadístico en Colombia. Su relevancia se respalda por el rol de la revista como un espacio que conecta académicos nacionales, consolidando un circuito interno de citas que fortalece la comunidad estadística del país.

A nivel internacional, la revista también muestra una conexión significativa con autores globalmente reconocidos, como Montgomery y Akaike, quienes aparecen constantemente citados en los trabajos publicados. Esto no solo denota la integración de teorías y métodos desarrollados internacionalmente, sino también cómo estas influencias externas moldean el marco conceptual y metodológico de los académicos locales. Herramientas como el lenguaje de programación R, frecuentemente citado en décadas recientes, ilustran cómo el conocimiento global ha influido en las prácticas estadísticas de Colombia, ampliando las capacidades técnicas y conceptuales de los investigadores.

Sin embargo, las métricas globales de centralidad como la de cercanía muestran que el impacto internacional de los académicos locales es relativamente limitado. Esto podría estar relacionado con la concentración de citas dentro de comunidades nacionales o regionales, lo que dificulta la difusión hacia redes más amplias y globales. Este fenómeno resalta la necesidad de estrategias que fortalezcan la proyección internacional, como fomentar colaboraciones interinstitucionales, publicar en revistas internacionales de mayor impacto o participar en congresos internacionales.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	Stock J H	F H Nieto	F H Nieto	Tsay R S	C G Mate
2	Tsay R S	Chen C W	Tsay R S	Stock J H	S Gallon
3	Saikkonen P	Tong H L	Tong H L	Phillips P B C	K Gomez
4	Phillips P B C	Tsay R S	Peña D	Watson M W	F H Nieto
5	Dickey D A	Peña D	Harvey A C	F H Nieto	E Castaño

Tabla 4: Top autores más importantes en la comunidad más grande de la red total

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.023361	0.310460	0.001023	61	193
2	0.007402	0.232845	0.000830	24	169
3	0.006534	0.232845	0.000829	23	164
4	0.006245	0.232845	0.000819	19	144
5	0.006245	0.232845	0.000812	19	119

Tabla 5: Valores de las medidas locales para clasificación del Top en la década de los 00s.

En conjunto, esta dinámica de citación pone a la vista un balance entre la con-

solidación de una comunidad estadística local robusta y el desafío de incrementar la visibilidad y contribución en el escenario internacional. La Revista Colombiana de Estadística, al actuar como un puente entre estas esferas, desempeña un papel esencial para impulsar el intercambio y la evolución del conocimiento estadístico tanto dentro como fuera del país.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.010691	0.780563	0.000391	191	1279
2	0.005976	0.314129	0.000349	114	447
3	0.004813	0.257015	0.000294	113	425
4	0.004368	0.190381	0.000276	75	354
5	0.004045	0.123748	0.000270	69	327

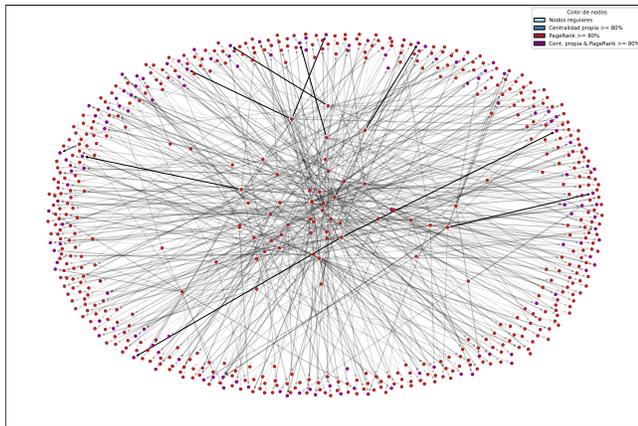
Tabla 6: Valores de las medidas locales para clasificación del Top en red total de citaciones entre autores.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.012152	0.486792	0.002508	20	221
2	0.011877	0.389434	0.002356	16	201
3	0.011392	0.243396	0.002183	14	201
4	0.010567	0.194717	0.002103	12	198
5	0.010155	0.194717	0.001898	12	123

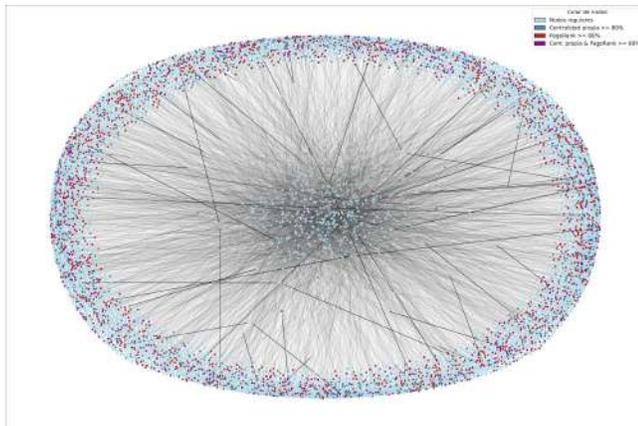
Tabla 7: Valores de las medidas locales para clasificación del Top en la comunidad de mayor tamaño de la red de autores.

4.2. Red de citaciones entre artículos

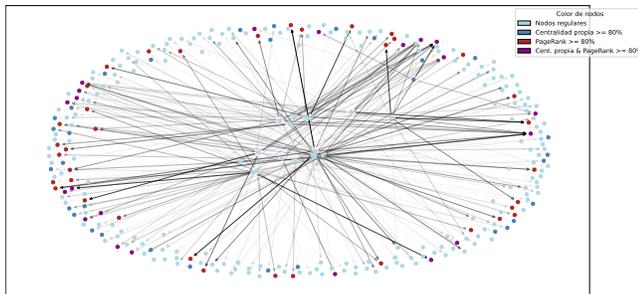
Teniendo en cuenta las representaciones de las redes de artículos, algunas de ellas mostradas en 5(b), se empieza analizando las décadas 60s y 70s. En la primera se destacan los artículos aislados que son considerados entre los más importantes dada su centralidad propia, lo cual puede suceder por los atributos que poseen estos nodos más que por sus relaciones internodales. Mientras que en la segunda década se evidencia un tamaño muy pequeño comparado al del resto de redes de años posteriores, razón por la cual se considera el fenómeno asociado al **factor de teletransportación** en el *PageRank*, este modela la probabilidad de que se referencie un artículo vs. la probabilidad de que "salte" aleatoriamente a otro artículo, implicando que hasta un nodo aislado pueda ser referenciado aleatoriamente, como sucede con el artículo "An introduction to linear statistical models". Al igual que en el tipo de redes anteriores, las redes de autores, los valores de centralidad de cercanía y *PageRank* adoptan valores muy pequeños a comparación de los obtenidos mediante la medida de centralidad propia, lo cual significa que los artículos clave en la conectividad de la red son clasificados por medio del número de artículos que los referencian y que son referenciados al mismo tiempo por un gran número de autores. Esto no impide comparar la importancia entre los artículos de las redes.



(a) Red de referencias entre los artículos publicados en la década de los 90s



(b) Red de referencias entre los artículos de la Revista Colombiana de Estadística



(c) Comunidad de mayor tamaño de la red de citas de artículos

Figura 5: Redes de artículos.

El contraste de tiempo entre publicación referenciada y publicación que referencia, es más notorio en la década de los 80s, teniendo las referencias hacia el artículo más importante de la red tienen una diferencia temporal de 426 años, mientras que en la década de los 00s los artículos que referencian a la publicación más importante de esta red tienen una diferencia temporal de 32 años, siendo más contemporáneas las referencias de esta década en particular. Estos datos se obtienen después de dividir la respectiva fuerza del artículo más importante por su grado de entrada. Respaldo el apogeo de referenciación en documentos computacionales de R que se observó anteriormente, esto mediante 9, 32, 34, se puede afirmar que desde los años 90s hasta la actualidad el artículo "R: A language and Environmental ..." juega un papel clave en las relaciones de citas, porque siempre se encuentra en el top 5 de artículos más importantes para dos, tres o hasta más medidas de centralidad. Al igual que el artículo "Linear models", un clásico presente desde los inicios de la revista que sigue siendo ampliamente citado en varias áreas del conocimiento.

Medida	60s	70s	80s	90s	2000	2010	2020	Completa	Comunidad
Num. Nodos	9	8	491	623	1841	4597	1962	9014	291
Num. Aristas	7	6	457	600	1930	4985	2005	9990	354
Grado Prom.	0.778	0.750	0.931	0.963	1.048	1.084	1.0219	1.108	1.216
Fuerza Prom.	187.111	13	63.866	39.226	35.719	38.064	37.744	41.311	41.484
Densidad	0.097	0.107	0.0019	0.0015	0.0006	0.0002	0.0005	0.0001	0.0042

Tabla 8: Medidas descriptivas globales de todas las redes de artículos.

Al realizar el particionamiento de la red total de citas entre artículos se obtuvieron 309 comunidades, de las cuales la de mayor tamaño poseía 291 nodos, los cuales representan el 3.22% de la totalidad de artículos, mientras que el tamaño promedio de las comunidades fue 30 y el tamaño mínimo 1, a pesar de tener artículos que se consideraban su propia comunidad, muy probablemente los nodos aislados encontrados en las décadas 60s y 70s, la **modularidad** de la partición dio un valor de 0.94, considerándolo como una partición casi perfecta. Al analizar la componente más grande, generada por la partición natural de la red, se evidencia que el nodo más importante "Consistent Estimates based on partially ..." es de los que tienen una menor centralidad propia dentro de los artículos más importantes de cada red, evidenciando que aunque en su comunidad es considerado el más clave, en la red total no juega un papel tan importante.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Detection Of One Upper Outlier In A Censored E...	A New Approach To Linear Filtering And Predict...	Convergence Of Probability Measures	Outliers In Statistical Data	Procesos Estocásticos Con Dos Parámetros 1
2	Bayesian Forecasting And Dynamic Models	The Identification Of Vector Mixed Autoregress...	El Censo De Soledad	NaN	Superficies De Respuesta Para Análisis De Dato...
3	Identificación De Un Modelo De Estados Para Un...	Markovian Representation Of Stochastic Process...	Nonparametric Statistical Inference	Bayesian Forecasting And Dynamic Models	Valoración De La Presión Del Neumotaponador En...
4	Análisis Estadístico De Fenómenos Atmosféricos...	Canonical Correlation Analysis Of Time Series ...	Queueing Systems With Vacation-A Survey	Convergence Of Probability Measures	Maximum-Likelihood Estimation Of The 4-Paramet...
5	Avances En El Conocimiento De Las Descargas El...	Sas/Ets User'S Guide	Una Prueba De Walel En Regresión	Outliers	Clasificación No Jerárquica Con Coordenadas Fa...

Tabla 9: Top artículos más importantes en la década de los 90s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	R: A Language And Environment For Statistical ...	A New Look At Statistical Model Identification	Sampling Techniques	R: A Language And Environment For Statistical ...	Fundamentos De La Teoría De Probabilides
2	Capability Indices For Birnbaum-Saunders Proce...	A Class Of Distributions Which Includes The No...	Life Contingencies	Continuous Univariate Distributions	Aplicaciones Del Análisis Factorial En Psicología
3	Slashed Rayleigh Distribution	Partially Adaptive Estimation Of Regression Mo...	Practical Nonparametric Statistics	Generalized Linear Models	A Methodology For Biplots Based On Bootstrappi...
4	Air Contaminant Statistical Distributions With...	Alternative Beta Estimation For The Market Mod...	Linear Models	Statistical Models And Methods For Lifetime Data	Bayesian Multi-Faceted Tri Models For Measurin...
5	A Generalized Class Of Ratio Type Exponential ...	Robust And Partially Adaptive Estimation Of Re...	The Analysis Of Variance	Linear Models	Relationship Between Kendall'S Tau Correlation...

Tabla 10: Top artículos más importantes de la Revista Colombiana de Estadística.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	A New Extension Of The Exponential Distribution	Consistent Estimates Based On Partially Consis...	A New Look At Statistical Model Identification	A Class Of Distributions Which Includes The No...	In Uence Diagnostics For Correlated Binomial R...
2	The Beta Generalized Rayleigh Distribution Wit...	Application Of The Radon-Nikodym Theorem To Th...	A Class Of Distributions Which Includes The No...	Estimating The Dimension Of A Model	Likelihood-Based Inference For The Asymmetric ...
3	Likelihood Based Inference For Distributions O...	The Unreplicated Ultrastructural Relation: Lar...	Estimating The Dimension Of A Model	Distributions Of Fractional Order Statistics L...	Robust Mixture Regression Based On The Skew T ...
4	Robust Mixture Regression Modeling Based On Sc...	Estimation In The Presence Of Incidental Param...	Distributions Of Fractional Order Statistics L...	A New Look At Statistical Model Identification	Properties And Inference For Proportional Haza...
5	Asymmetric Regression Models With Limited Resp...	A Formula For The Distribution Of The Maxim...	Data: A Collection Of Problems From Many Field...	A Probabilistic Representation Of The Skew-Nor...	An Empirical Comparison Of Em Initialization M...

Tabla 11: Top artículos más importantes la comunidad de mayor tamaño de la red de citaciones de artículos.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.003215	0.701287	0.003025	4	970.0
2	0.002572	0.490901	0.002727	4	958.0
3	0.002144	0.374020	0.002648	3	636.0
4	0.001608	0.327268	0.002383	3	622.0
5	0.001608	0.140258	0.002341	3	599.0

Tabla 12: Valores de las medidas locales para clasificación del Top en la década de los 90s.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.047619	0.693805	0.163759	1	26.0
2	0.035714	0.536655	0.146907	1	26.0
3	0.023810	0.320218	0.130055	1	18.0

Tabla 13: Valores de las medidas locales para clasificación del Top en red total de citaciones entre autores.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.012152	0.486792	0.002508	20	221
2	0.011877	0.389434	0.002356	16	201
3	0.011392	0.243396	0.002183	14	201
4	0.010567	0.194717	0.002103	12	198
5	0.010155	0.194717	0.001898	12	123

Tabla 14: Valores de las medidas locales para clasificación del Top en la comunidad de mayor tamaño de la red de autores

5. Conclusiones

El análisis de las redes de citaciones en la Revista Colombiana de Estadística permite identificar patrones significativos en la producción académica en el ámbito de la estadística. A nivel nacional, los docentes de la Universidad Nacional de Colombia se destacan como actores clave en la divulgación y consolidación del conocimiento estadístico. Su influencia se refleja en su posición central dentro de las redes analizadas, donde actúan como nodos importantes en la estructura de colaboración académica en Colombia. Esto refuerza el papel de la Universidad Nacional como una institución que no solo forma profesionales en estadística, sino que también impulsa la investigación en este campo. Mientras que a nivel global, el análisis muestra la creciente relevancia de herramientas computacionales, como el lenguaje R, que se han convertido en elementos indispensables para la investigación estadística moderna. Estas herramientas no solo han facilitado la implementación de métodos estadísticos, sino que también han generado un impacto significativo en las relaciones de citaciones, marcando un cambio en la forma en que se difunde y se aplica el conocimiento en estadística.

Por otro lado, se observa una evolución en las dinámicas de citación a lo largo de las décadas. Mientras que las primeras décadas estuvieron marcadas por un enfoque en autores pioneros y métodos clásicos, las décadas más recientes han estado dominadas por la integración de metodologías computacionales y la aplicación de técnicas estadísticas avanzadas. Lo cual refleja un cambio en las prioridades de la comunidad académica y en la manera en que la estadística se posiciona como disciplina en un entorno cada vez más tecnológico e interconectado. Pero a pesar de los avances, el análisis también evidencia ciertas limitaciones en la proyección internacional de los autores nacionales, destacando la necesidad de fortalecer las estrategias de colaboración internacional y de publicar en revistas de mayor impacto global. Sin embargo, el papel de la revista como puente entre las comunidades académicas nacionales e internacionales es innegable y sigue siendo un pilar en la evolución de la estadística en Colombia.

En conjunto, los resultados del estudio no solo proporcionan una visión detallada de las interacciones académicas en el contexto Colombiano, sino que también resaltan la importancia de las redes de citaciones como herramienta para comprender la evolución del conocimiento en estadística y la influencia de las comunidades académicas en su desarrollo.

6. Limitaciones y trabajo futuro

El estudio presenta limitaciones significativas, principalmente relacionadas con la calidad y completitud de los datos de décadas tempranas, lo que dificulta la obtención de conclusiones generales. La baja densidad de las redes también afecta la robustez de las métricas utilizadas, como la centralidad de cercanía y PageRank, que pueden subestimar el impacto global de ciertos nodos. Además, al enfocar-

se exclusivamente en la Revista Colombiana de Estadística, los hallazgos no son extendibles a otras comunidades académicas.

En investigaciones futuras, sería valioso ampliar el análisis a datos más diversos, considerando múltiples revistas y contextos internacionales para obtener una visión más integral. Incorporar métricas más robustas, como la centralidad de Katz, permitirá capturar mejor las dinámicas en redes con baja densidad. También sería interesante explorar el impacto de colaboraciones internacionales y la creciente importancia de herramientas computacionales en la generación de conocimiento estadístico.

Recibido:

Aceptado:

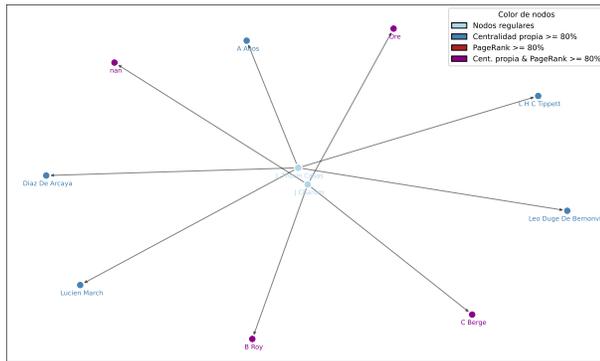
Referencias

- S. Bianco. *Revista Latinoamericana de Ingeniería de Software*, 4(2):111, 2016. doi: 10.18294/relais.2016.111-142.
- V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008. doi: 10.1088/1742-5468/2008/10/p10008.
- D. Maltseva and V. Batagelj. Journals publishing social network analysis. *Scientometrics*, 126(4):3593–3620, 2021. doi: 10.1007/s11192-021-03889-z.
- F. E. Muñoz R. Material de apoyo unidad 4 estructura de datos, 2015. URL <https://slideplayer.es/slide/3356703/>.
- NetworkX. Centrality, 2023. URL <https://networkx.org/documentation/stable/reference/algorithms/centrality.html>.
- E. Otte and R. Rousseau. Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, 28(6):441–453, 2002. doi: 10.1177/016555150202800601. URL <https://doi.org/10.1177/016555150202800601>.
- L. Page. Method for node ranking in a linked database. united states patent, 2001. URL <https://patentimages.storage.googleapis.com/37/a9/18/d7c46ea42c4b05/US6285999.pdf>.
- J. C. Sosa. Redes sociales, 2022. URL <https://sites.google.com/view/juansosa/social-networks>.
- Wikipedia. Grafo completo, 2020. URL https://es.wikipedia.org/wiki/Grafo_completo.

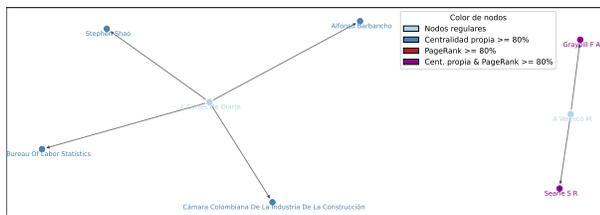
A. Anexos

A.1. Redes de autores

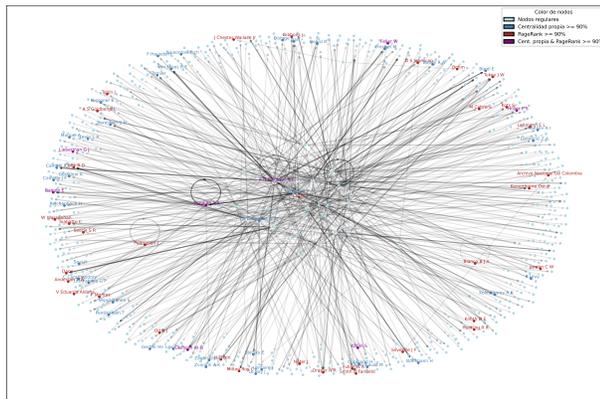
A.1.1. Grafos



(a) Red de citas entre autores para artículos publicados en la década de los 60s

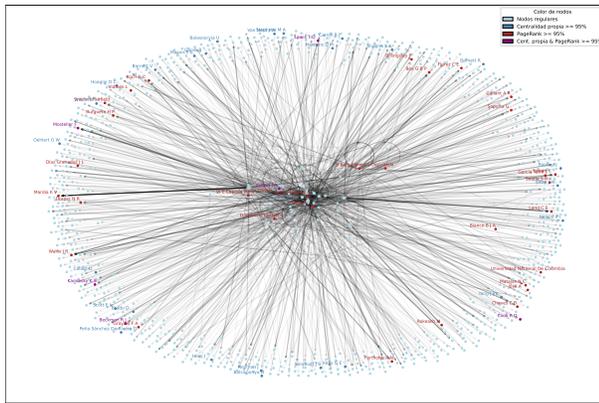


(b) Red de citas entre autores para artículos publicados en la década de los 70s

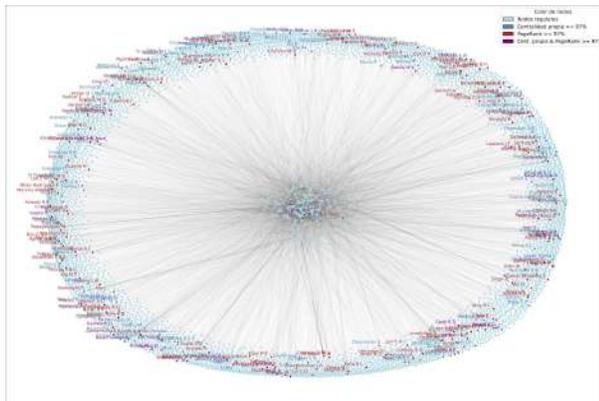


(c) Red de citas entre autores para artículos publicados en la década de los 80s

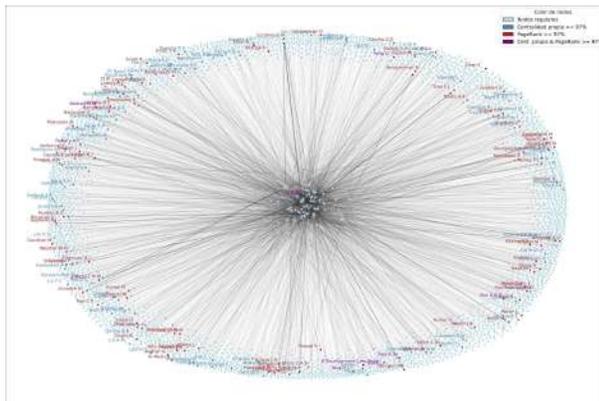
Figura 6: Redes de citas entre autores para las décadas 60s-80s.



(a) Red de citaciones entre autores para artículos publicados en la década de los 90s



(b) Red de citaciones entre autores para artículos publicados en la década de los 2010



(c) Red de citaciones entre autores para artículos publicados en la década de los 2020

Figura 7: Redes de citaciones entre autores para las décadas 90s-2020 exceptuando la década de los 2000.

A.1.2. Tablas de clasificación de autores más importantes en redes de autores

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Todos los autores	Todos los autores	B Roy	Todos los autores	L Thorin Casas
2	menos L Thorin Casas	menos L Thorin Casas	NaN	menos L Thorin Casas	J Charum
3	y J Charum	y J Charum	C Berge / Ore	y J Charum	

Tabla 15: Top autores más importantes en la década de los 60s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Todos los autores menos	Todos los autores menos	Searle S R	Todos los autores menos	L Cortes De Olarte
2	L A Cortes De Olarte M	L A Cortes De Olarte	Graybill F A	L A Cortes De Olarte	A Velasco M
3	y A Velasco	y A Velasco	Cámara Colombiana De La Industria De La Construcción		

Tabla 16: Top autores más importantes en la década de los 70s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Smith H Fairfield	Fandiño A R	Poltronieri J	Smith H Fairfield	Ospina D
2	Cochran W G	Ramón E	M Cabrera	Tukey J W	Cepeda C F
3	Kemphorne Oscar	Cepeda C F	Tobín J	Draper N R	E Castaño V
4	Tukey J W	A N Kolmogorov	F Montes	Cochran W G	A N Kolmogorov
5	Lehmann E L	Maltsev A I	Dane	Kemphorne Oscar	R Matos Marengo

Tabla 17: Top autores más importantes en la década de los 80s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Smith H Fairfield	O Barbary	Mardia K V	Smith H Fairfield	P N Pacheco D
2	Box G B P	Saporta G	Wallis J R	Box G B P	Bautista L
3	Mosteller F	Florez C E	Universidad Nacional De Colombia	Mosteller F	M Muñoz De Ozak
4	Graybill F A	K C Madan	O Barbary	Sas Institute	F H Nieto
5	Stephens M A	Deville J C	Blanco B J A	Amiltage P	M Garzon

Tabla 18: Top autores más importantes en la década de los 90s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	R Development Core Team	V Leiva	Balakrishnan N	R Development Core Team	V Leiva
2	Kotz S	Sanhueza A	R Development Core Team	Balakrishnan N	M P Galindo
3	Balakrishnan N	Balakrishnan N	Singh H P	Kotz S	J A Achcar
4	Kaplan E L	Galindo M P	Kotz S	Kundu D	G Martinez Florez
5	Tibshirani R J	Marchant C	V Leiva	Nadarajah S	O Arslan

Tabla 19: Top autores más importantes en la década de los 2010 según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	R Development Core Team	F H Nieto	Cordeiro G M	R Development Core Team	B Oluyede
2	Rubin D B	Chen C W	Balakrishnan N	Nadarajah S	J C Sosa
3	Kotz S	Tong H L	Al-Masri A	Cordeiro G M	M Ndiaye
4	Nadarajah S	Tsay R S	Nadarajah S	Balakrishnan N	S Dabo Niang
5	Lawless J F	Peña D	Al Zahrani B M	Rubin D B	P Ngom

Tabla 20: Top autores más importantes en la década de los 2020 según diferentes medidas locales.

Tabla 21: Top autores más importantes en las década de los 60s-2020 exceptuando la década de los 2000.

A.1.3. Medidas locales para el Top 3 Autores

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.1	0.333333	0.095472	1	5
2	0.1	0.333333	0.095472	1	4
3	0.1	0.333333	0.095472	1	

Tabla 22: Valores de las medidas locales para clasificación del Top en la década de los 60s.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.142857	0.408248	0.146907	1	4
2	0.142857	0.408248	0.146907	1	2
3	0.142857	0.408248	0.125000	1	

Tabla 23: Valores de las medidas locales para clasificación del Top en la década de los 70s.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.011194	0.689962	0.002934	6	49
2	0.010855	0.413977	0.002676	6	49
3	0.009328	0.275985	0.002676	6	48
4	0.009142	0.206993	0.002676	6	45
5	0.008396	0.137992	0.002660	5	35

Tabla 24: Valores de las medidas locales para clasificación del Top en la década de los 80s.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.010753	0.564329	0.001915	8	95
2	0.008233	0.338598	0.001820	7	69
3	0.007467	0.338598	0.001741	7	63
4	0.007258	0.282165	0.001729	5	61
5	0.007168	0.225732	0.001727	5	53

Tabla 25: Valores de las medidas locales para clasificación del Top en la década de los 90s.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.008232	0.780634	0.000533	116	1254
2	0.005891	0.314158	0.000477	90	354
3	0.004724	0.257038	0.000453	86	342
4	0.004346	0.190398	0.000420	48	298
5	0.004065	0.123759	0.000416	47	285

Tabla 26: Valores de las medidas locales para clasificación del Top en la década de los 2010.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.008982	0.502513	0.000548	29	425
2	0.005204	0.402011	0.000484	28	232
3	0.005188	0.251257	0.000460	27	200
4	0.004430	0.201005	0.000433	25	200
5	0.004272	0.201005	0.000430	18	200

Tabla 27: Valores de las medidas locales para clasificación del Top en la década de los 2020

Tabla 28: Valores de medidas locales para clasificación del Top autores más importantes en la década de los 60s-80s.

A.2. Redes de artículos

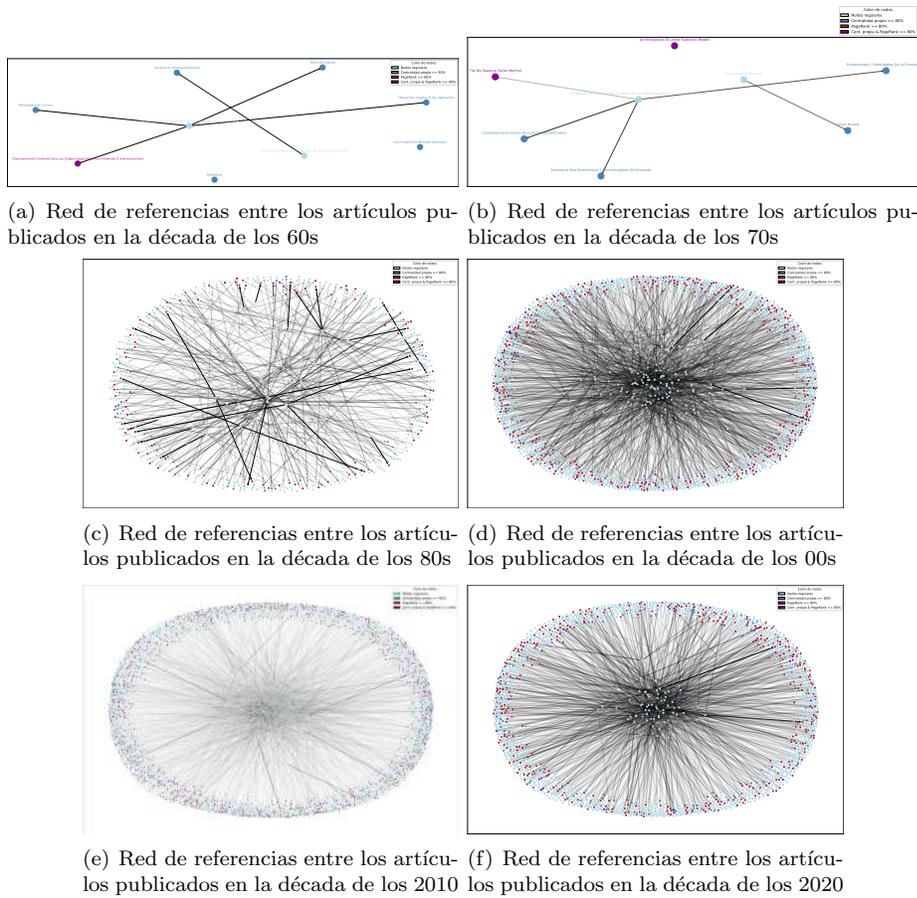


Figura 8: Redes de citaciones entre artículos para las décadas 60s-2020 excepto la década de los 00s.

A.2.1. Tablas de clasificación de artículos más importantes en redes de artículos

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Teoría De Gráficos	Les Principes Del Methode Statistique	Cheminement Et Connexité Dans Les Graphes Appl...	Iniciatian A L'Analyse Statistique	La Estadística Elemental Como Parte De La Educ...
2	Policopiados De La Ensa	Estadística	Estadística	Estadística	Estadística
3	Théorie Des Graphes Et Ses Applications	Iniciatian A L'Analyse Statistique	Les Principes Del Methode Statistique	Cheminement Et Connexité Dans Les Graphes Appl...	Les Principes Del Methode Statistique

Tabla 29: Top artículos más importantes en la década de los 60s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Comportamiento Cíclico De La Actividad Edifica...	The Bls Seasonal Factor Method	An Introduction To Linear Statistical Models	Fundamentos Y Posibilidades De La Econometría	El Método Bls Para La Determinación De La Esta...
2	Estadística Para Economistas Y Administradores...	Fundamentos Y Posibilidades De La Econometría	The Bls Seasonal Factor Method	An Introduction To Linear Statistical Models	La Inversa De Penrose
3	damentos Y Posibilidades De La Econometría	Estadística Para Economistas Y Administradores...	Linear Models	Linear Models	An Introduction To Linear Statistical Models

Tabla 30: Top artículos más importantes en la década de los 70s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	Matemática Financiera, Enfoque Toma De Decisiones	Múltiple Frame Surveys	Life Contingencies	Applied Regression Analysis	Fundamentos De La Teoría De Probabilides
2	Reflexión De Matrices Y Palíndromía	Sainplng Tech	he Analysis Of V_c	Design And Analysis Of Experiments	Aplicaciones Del Análisis Factorial En Psicología
3	Experimental Design: Procedures For The Behavi...	Theory And Application Of Multiple Frame Surveys	Practical Nonparametric Statistics	NaN	Una Distribución Asintótica Para Estimadores M...
4	Procesos Markovianos Del Sistema De Descuentos...	Monetary Policy Inflation And Unemployment	A First Course In Stochastic Processes	Life Contingencies	Enseñanza Y Currículo De Estadística En Difere...
5	Contribución Al Estudio De La Actividad Sísmic...	La Dinámica De Las Series De Tiempo: El Cambio...	Matemática Financiera, Enfoque Toma De Decisiones	Econometric Methods	Contribución Al Estudio De Formas Cuadráticas ...

Tabla 31: Top artículos más importantes en la década de los 80s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	R: A Language And Environment For Statistical	Calculating The Singular Values And Pseudoinve...	Linear Models	R: A Language And Environment For Statistical ...	Muestreo De Conglomerados Con Multiplicidad: E...
2	Weaker Conditions For Asymptotic Approximation...	Time Series Analysis, Forecasting And Control	Sampling Techniques	Linear Models	Optimización De Curvas De Crecimiento A Través...
3	Criterios De Selección Y UtilizacióN De Info...	Distribution Of Residual Autocorrelations In A...	Model Assisted Survey Sampling	Generalized Linear Models	Comparación Entre Tres Técnicas De Clasificación
4	K-Sample Tests Based On The Likelihood Ratio	On A Measure Of Lack Of Fit In Time Serie Moldels	RecolectióN Y Análisis De Datos Longitudinales	Model Assisted Survey Sampling	Tendencia Aleatoria O Determinística: Una Nuev...
5	Shared Random Effects Analysis Of Multi-State ...	On The Distribution And Application Of Residua...	Two Rules Of Thumb For The Approximation Of Th...	Statistical Analysis With Missing Data	Modelado De Parejas Aleatorias Usando Cópulas

Tabla 32: Top artículos más importantes en la década de los 00s según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	R: A Language And Environment For Statistical ...	A New Look At The Statistical Model Identifi CCation	The Poisson-Lomax Distribution	R: A Language And Environment For Statistical ...	Bayesian Multi-Faceted Tri Models For Measurin...
2	The Weibull Marshall-Olkin Lindley Distributio...	Equation Of State Calculations By Fast Computi...	A Mathematical Theory of Communication	Stochastic Orders	Relationship Between Kendall'S Tau Correlation...
3	Reparameterized Inverse Gamma Regression Model...	Estimating The Dimension Of A Model	Stochastic Comparison Of Tests	Statistical Methods For Reliability Data	A Review Of Latent Space Models For Social Net...
4	Testing The Degree Of Overlap For The Expected...	Algorithm As 217: Computation Of The Dip Stati...	Less Vulnerable Con Dence And Significance Proc...	Statistical Models And Methods For Lifetime Data	Variable Selection In Switching Dynamic Regres...
5	Estimation Of Stress-Strength Reliability Usin...	A Class Of Distributions Which Includes The No...	X - Contributions To The Mathematical Theory O...	Estimating The Dimension Of A Model	Optimal Detection Of Bilinear Dependence In Sh...

Tabla 33: Top artículos más importantes en la década de los 2020 según diferentes medidas locales.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	R: A Language And Environment For Statistical ...	A New Look At Statistical Model Identification	A New Look At Statistical Model Identification	R: A Language And Environment For Statistical ...	A Methodology For Biplots Based On Bootstrappi...
2	Capability Indices For Birnbaum-Saunders Proce...	A Class Of Distributions Which Includes The No...	A Class Of Distributions Which Includes The No...	Continuous Univariate Distributions	Bayesian Analysis Of The 3-Component Mixture O...
3	Slashed Rayleigh Distribution	Partially Adaptive Estimation Of Regression Mo...	Nonparametric Estimation From Incomplete Obser...	A Class Of Distributions Which Includes The No...	Determinantes Socioeconómicos De La Mortalidad...
4	Air Contaminant Statistical Distributions With...	Alternative Beta Estimation For The Market Mod...	Sampling Theory And Methods	An Introduction To Copulas	Statistical Properties And Different Methods O...
5	A Generalized Class Of Ratio Type Exponential ...	Robust And Partially Adaptive Estimation Of Re...	Bivariate Exponential Distributions	Statistical Models And Methods For Lifetime Data	Kernel Function In Local Linear Peters-Belson ...

Tabla 34: Top artículos más importantes en la década de los 2010 según diferentes medidas locales.

A.2.2. Medidas locales para el Top 3 en las redes de artículos

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.0625	0.705777	0.136233	1	829.0
2	0.0625	0.705777	0.131788	1	400.0
3	0.0625	0.051169	0.131788	1	400.0

Tabla 35: Valores de las medidas locales para clasificación del Top en la década de los 60s

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.047619	0.693805	0.163759	1	26.0
2	0.035714	0.536655	0.146907	1	26.0
3	0.023810	0.320218	0.130055	1	18.0

Tabla 36: Valores de las medidas locales para clasificación del Top en la década de los 70s.

Top	Cent. Cercanía	Cent. Propia	PageRank	Grado de entrada	Fuerza
1	0.002721	0.961976	0.004989	5	2134.0
2	0.002721	0.154441	0.003946	3	1525.0
3	0.002041	0.146310	0.003625	3	1048.0
4	0.002041	0.130056	0.003588	3	880.0
5	0.002041	0.064933	0.003388	2	843.0

Tabla 37: Valores de las medidas locales para clasificación del Top en la década de los 80s.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.013443	0.466866	0.001094	28	830.0
2	0.002174	0.405437	0.001093	9	819.0
3	0.001087	0.405437	0.000978	7	786.0
4	0.001087	0.307149	0.000874	7	740.0
5	0.001087	0.307149	0.000836	6	679.0

Tabla 38: Valores de las medidas locales para clasificación del Top en la década de los 00s.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.004528	0.427669	0.000399	47	1479.0
2	0.001958	0.307044	0.000352	14	1167.0
3	0.000979	0.296078	0.000350	10	1105.0
4	0.000979	0.285112	0.000337	9	1099.0
5	0.000979	0.274147	0.000322	8	998.0

Tabla 39: Valores de las medidas locales para clasificación del Top en la década de los 2010.

Top	Cent. Cercanía	Cent. Propia	<i>PageRank</i>	Grado de entrada	Fuerza
1	0.001492	0.508456	0.000916	12	1207.0
2	0.000656	0.450529	0.000780	5	1194.0
3	0.000510	0.386170	0.000683	4	1153.0
4	0.000510	0.296063	0.000663	4	1046.0
5	0.000510	0.263882	0.000660	4	1045.0

Tabla 40: Valores de las medidas locales para clasificación del Top en la década de los 2020.